

# Multi-outcome and Multidimensional Market Scoring Rules (Manuscript)

Lance Fortnow\*

Northwestern University

*fortnow@eecs.northwestern.edu*

Rahul Sami†

University of Michigan

*rsami@umich.edu*

February 9, 2012

## Abstract

Hanson’s market scoring rules allow us to design a prediction market that still gives useful information even if we have an illiquid market with a limited number of budget-constrained agents. Each agent can “move” the current price of a market towards their prediction.

While this movement still occurs in multi-outcome or multidimensional markets we show that no market-scoring rule, under reasonable conditions, always moves the price directly towards beliefs of the agents. We present a modified version of a market scoring rule for budget-limited traders, and show that it does have the property that, from any starting position, optimal trade by a budget-limited trader will result in the market being moved towards the trader’s true belief. This mechanism also retains several attractive strategic properties of the market scoring rule.

## 1 Introduction

Prediction markets are markets set up for the primary purpose of aggregating information to forecast future events. For a heavily traded prediction market, such as whether Obama will be reelected in 2012, we can get a very precise prediction from the market. As of January 30, 2012, the Intrade bid-ask spread gives a prediction between 53.3% and 53.9% of reelection. For other less liquid markets, the bid-ask spread gives a less informative prediction. The Intrade bid-ask spread for whether Sarah Palin will endorse Mitt Romney is sitting between 15.1% and 99.0%.

To recover information in thin, illiquid markets, Robin Hanson developed the market scoring rule [Han03]. With this form prediction market, a market maker always quotes a single current price for each future events. If the agents have a belief different than the current price, a trade will give them an a priori expected positive

---

\*Supported in part by NSF grant CCF-0829754

†Supported in part by NSF grant CCF-0812042

profit. The market maker adjusts the price as trades are made. If traders are risk-neutral and do not face budget constraints, Hanson showed that it is optimal for a trader to trade until the price matches her beliefs; the final market prices can thus be interpreted as updated probabilities for the forecast events. The market maker loses money in expectation, but the total loss of the market maker can be bounded.

The further the change in price, the higher the up-front cost, and worst-case loss, that the trader has to bear. In real-world prediction markets, traders often have budget constraints that limit how much loss they can bear. In fact, many real money prediction market platforms impose limits on how much money an agent can invest over the entire platform, and most play-money markets limit the ability of traders to convert external money to virtual points that can be invested. Thus, limited budgets are an inherent feature of the context in which most prediction markets are run. If an agent is budget-limited she can still trade, but may not be able to move the prices to match her beliefs. For two-outcome markets, like the ones described above, the market scoring rule induces straightforward behavior in budget-limited agents: An agent will move the prices of the securities directly towards her belief, and the final probability forecast will therefore be a mixture of the original market probabilities and the agent's true belief. In this paper, we explore the behavior of budget-limited traders in markets with multiple outcomes or multi-dimensional outcomes.

Many natural events for which we seek forecasts have multiple outcomes. Consider the following scenarios:

- A market over Republican candidates for the U.S. Presidential nomination: For this market, a forecast can be represented by a four-tuple probability vector  $(p_R, p_G, p_P, p_S)$  for the probabilities that Romney, Gingrich, Paul and Santorum will be nominated.
- A market predicting the latitude and longitude of the next tornado strike in Kansas. Here, the forecasts consist of probability distributions over a two-dimensional grid.
- A market predicting the likelihood of hurricane strikes for each year over the next three years. Here, a forecast consists of a probability distribution over  $\{0, 1\}^3$ ; however, it may be natural to assume independence between different years.

For each of these markets, we can set up prediction markets based on proper scoring rules. One way to do so is to break the markets into their components, a separate market for each candidate, for latitude and longitude and for each year. More generally, however, we may have a complex combinatorial market whose payoffs are dependent on the combination of events that occur [Han03]. In either case, a risk-neutral trader without budget constraints would optimize her payoffs by moving the market forecast to match her true belief. When traders are budget-limited, the optimal behavior may depend on the particular proper scoring rule used. One natural goal is to ask: Can we find a proper scoring rule such that, for any positive budget an agent has, her optimal trade is always to move directly towards her true belief, so that the updated forecast is a mixture of the original forecast distribution and her true belief?

Our main results show that this is impossible. In section 3 we show that, under mild smoothness assumptions, no scoring rule over three or more outcomes can have this property. For example, in the Republican candidate setting, we cannot find any market scoring rule that will induce a budget-constrained trader to move the market forecast in a straight line towards her true belief. Further, we show that there are neighborhoods of possible beliefs that all result in exactly the same report; this shows that sequential aggregation in the MSR with budget constraints can be imperfect, because even if all budgets are public information, a trader’s belief cannot be inferred from her report. The results in section 3 relies on the space of beliefs being unrestricted, and thus seems to leave open the possibility of a positive result when there are natural restrictions, such as independence of hurricane probabilities in different years. However, we show that this is not the case in section 4: Even when the outcome space consists of two binary events that are assumed to be independent, no smooth market scoring rule can induce a budget-limited agent to move the market distribution directly towards her true belief. Importantly, this latter result holds without assuming that the scoring rule is itself simple or separable, i.e., it may be a combinatorial scoring rule.

In section 5, we present a modified market scoring rule, the Scaled Scoring Mechanism, that has this property of inducing budget-limited agents to move the forecast directly towards their true belief. This mechanism requires a trader to report her budget as well as her belief; we assume that the budget must be deposited up-front, and hence cannot be exaggerated (although it may be under-reported). We prove that a trader’s optimal trade is to report her true budget and belief, and hence, the updated market forecast is always a mixture of her true belief and the original forecast. Further, we show that this mechanism retains many of the desirable properties of the market scoring rule: The market-maker’s loss is bounded, and the traders’ profit along a sequence of updates is sub-additive, so that a trader cannot gain a profit from splitting up her trade into two or more successive trades by different identities. As this mechanism induces full information revelation from budget-limited traders, it also enables effective sequential aggregation of information: A trader who has seen previous traders’ trades can infer their beliefs; hence, in a Bayesian world, a trader can generically condition her beliefs on all previous traders’ private information.

The rest of this paper is structured as follows: We discuss related work in section 1.1. We introduce our formal model in section 2. We prove the impossibility results in section 3 and 4, and introduce the new mechanism in section 5. Finally, in section 6, we discuss extensions and limitations of our results, and directions for future work.

## 1.1 Related Work

There is a long literature on the use of proper scoring rules to incentivize accurate probabilistic forecasts of events [Goo52] [WM68]. Proper scoring rules are characterized by the property that, given any belief  $\mathbf{p}$ , the report that globally maximizes the forecaster’s expected (under belief  $\mathbf{p}$ ) reward is the honest report  $\mathbf{q} = \mathbf{p}$ . Subsequent research on *effective scoring rules* studied the design of scoring rules with a stronger monotonicity property. A scoring rule is said to be effective with respect to a given metric  $L$  over the space

of probability distributions if, given any two possible reports  $\mathbf{q}, \mathbf{q}'$  such that  $\mathbf{q}$  is closer (under metric  $L$ ) to the true belief  $\mathbf{p}'$  than  $\mathbf{q}'$ , the expected score of reporting  $\mathbf{q}$  must be greater than the expected score of reporting  $\mathbf{q}'$  [Fri83]. Friedman [Fri83] has shown that there are scoring rules that are effective with respect to some metrics, including the  $L_2$  metric. Our negative results present a striking contrast: We show that no scoring rule can induce a budget-limited agent to move only along the direction of steepest decrease in  $L_2$  distance to the true belief.

Hanson’s market scoring rules [Han03] transform scoring rules into a mechanism for obtaining sequential forecasts from multiple forecasters. Hanson showed that this sequential scoring form can serve as an automated market maker for prediction markets.

Budget limits can be viewed as a special form of risk aversion. Allen [All87] described a lottery technique to extend scoring rules to forecasters with arbitrary and unknown risk aversion. Dimitrov et al. [DSE09] show that this technique can be extended to a sequential forecasting mechanism. However, the mechanism of Dimitrov et al. has the undesirable property that expected profits shrink exponentially as the number of traders grows, and moreover, they prove that this is unavoidable in the context of mechanisms for arbitrarily risk averse agents. The mechanism we present here does not have this undesirable property.

Manski [Man04] and Wolfers and Zitzewitz [WZ06] have also studied the effect of budget limits on the aggregative properties of prediction markets. However, our model differs from their models in significant ways: They study a stylized unsubsidized market model in which clearing prices are determined in a one-shot game, whereas we model a market scoring rule with common0knowledge Bayesians who update beliefs after each trade.

## 2 Model

**Outcomes and Distributions:** Consider an event with more than two possible outcomes. In order to simplify the notation, we begin with a three-outcome event, and let the outcomes be denoted  $X, Y, Z$ . An agent has a belief  $\mathbf{p} = (p_X, p_Y, p_Z)$  on the probabilities of the various outcomes. When asked for a forecast, the agent may report a forecast  $\mathbf{q} = (q_X, q_Y, q_Z)$ . In general, for unrestricted beliefs over a  $k$ -outcome event,  $\mathbf{p}$  and  $\mathbf{q}$  can be represented as  $k$ -dimensional vectors, with the implicit constraints that the elements are non-negative and sum to 1.

We also consider a special, natural, class of *restricted* distributions, which arise when the event being forecast itself has a dimensional structure. For clarity, we defer the description of this restricted class to section 4.

**Scoring Rules** The agent is rewarded based on a scoring rule  $\mathbf{S}$ . We model  $\mathbf{S}$  as a mapping from the space of forecasts  $\mathbf{q}$  to the space of payments  $\mathbf{S}(\mathbf{q}) \equiv (S_X(\mathbf{q}), S_Y(\mathbf{q}), S_Z(\mathbf{q}))$  that will be made for each outcome.

For unrestricted beliefs over three-outcome events, the space of all probability distributions  $\Delta = \Delta(\{X, Y, Z\})$  is a two-dimensional space. The scoring rule  $\mathbf{S}$  can be viewed as a two-dimensional surface in  $\Re^3$ ; each point  $\mathbf{S}(\mathbf{q}) \in \Re^3$  corresponds to a particular value of  $\mathbf{q}$ .

**Assumptions:** We assume that the scoring rule  $\mathbf{S}(\mathbf{q})$  is continuous and differentiable over the interior of the space of possible reports  $\mathbf{q}$ . (Actually, all we really need is that it is continuous and differentiable over some open set of reports.) We also assume that each component of the scoring rule is quasiconcave:  $\forall \mathbf{p}, \mathbf{q} \ S_X(0.5\mathbf{p} + 0.5\mathbf{q}) \geq \min(S_X(\mathbf{p}), S_X(\mathbf{q}))$ , with equality only if  $S_X(\mathbf{p}) = S_X(\mathbf{q})$ ; and likewise for  $S_Y, S_Z$ . This is a mild condition (weaker than monotonicity of the score functions), that is true for all commonly used scoring rules. It is used to argue that, when moving from  $\mathbf{q}$  to  $\mathbf{p}$ , the budget increases monotonically along the path, and so any point on the path is optimal for some budget. We assume that agents are risk neutral (up to their hard budget constraints, described below); thus, an agent's goal is to maximize her expected score  $\mathbf{p} \cdot \mathbf{S}(\mathbf{q})$ .

**Market Scoring Rules and Sequential Updates** Hanson's market scoring rules (MSR) [Han03] are a form of prediction market that extends scoring rules to a sequential information aggregation mechanism. Consider a set of agents  $\{1, 2, \dots, m\}$  making sequential forecasts  $\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_m$ . With the market scoring rule based on scoring rule  $\mathbf{S}$ , each agent  $i$  receives a net payoff equal to the difference between the score of  $\mathbf{q}_i$  and  $\mathbf{q}_{i-1}$ . Thus, the net expected payoff of an agent  $i$  who believes a distribution  $\mathbf{p}$  is:

$$\mathbf{p} \cdot [\mathbf{S}(\mathbf{q}_i) - \mathbf{S}(\mathbf{q}_{i-1})]$$

Here,  $\mathbf{q}_0$  is an initial default distribution specified by the mechanism.

Hanson observed that, with the MSR mechanism, if each agent trades just once and there are no budget constraints, the optimal action for each trader is to report  $\mathbf{q}_i$  that matches her true belief. Consequently, future traders who can observe the past reports (prices) in the market can condition on all previous agents' beliefs about the event. Generically, this leads to perfect aggregation of the  $m$  agents' combined information.

In our analysis of scoring with budget constraints, each agent's report  $\mathbf{q}$  is scored with respect to a *reference point*  $\bar{\mathbf{q}}$ , and thus her score is given by the difference  $\mathbf{S}(\mathbf{q}) - \mathbf{S}(\bar{\mathbf{q}})$ . The notion of a reference point or default  $\bar{\mathbf{q}}$  is an essential ingredient of a model with budget constraints, because it describes what the forecast will be when an agent has 0 budget. The reference point may be the previous traders' report  $\mathbf{q}_{i-1}$ , as in the market scoring rule, but it may also be determined differently; in section 5 we present a mechanism with a different way of determining a reference point for each trader.

**Budgets and Constrained Effectiveness** We are interested in scoring rules that are strictly proper: The unique optimal  $\mathbf{q}$ , given belief  $\mathbf{p}$ , is  $\mathbf{q} \equiv \mathbf{p}$ . However, we are further interested in a *budget-limited* version of this property. Informally, we want the scoring rule to have the property that, given any initial value  $\bar{\mathbf{q}}$ , and any specified budget  $b$ , the optimal report within the budget feasible set is a report  $\mathbf{q}$  along the path from  $\bar{\mathbf{q}}$  to  $\mathbf{p}$ .

In order to make this more rigorous, we need to define the notion of a budget more formally. One natural definition of the budget required to move from  $\bar{\mathbf{q}}$  to  $\mathbf{q}$  is in terms of the maximum loss that the agent can incur from such a move; this is the amount of money that an agent would require to have to avoid defaulting on the mechanism.

**Definition 1** *The natural budget  $nb(\bar{\mathbf{q}}, \mathbf{q})$  required to move from  $\bar{\mathbf{q}}$  to  $\mathbf{q}$ , given a scoring rule  $\mathbf{S}$ , is defined by*

$$nb(\bar{\mathbf{q}}, \mathbf{q}) \stackrel{def}{=} \max_{\mathbf{p}} [\mathbf{p} \cdot (\mathbf{S}(\bar{\mathbf{q}}) - \mathbf{S}(\mathbf{q}))] = \max\{S_X(\bar{\mathbf{q}}) - S_X(\mathbf{q}), S_Y(\bar{\mathbf{q}}) - S_Y(\mathbf{q}), S_Z(\bar{\mathbf{q}}) - S_Z(\mathbf{q})\}$$

Together, the reference point  $\bar{\mathbf{q}}$  and an initial budget holding  $b$  determine the range of possible reports that an agent can make while avoiding default under any outcome; the natural budget constraint determines the set of feasible reports. Thus, the natural budget constraint determines a budget-constrained agents' choice set under the market scoring rule as typically implemented. However, it may be possible for an alternative mechanism to further restrict the set of allowable reports by an agent; in section 5, we will exploit this to obtain a mechanism with better aggregative properties.

Now, the *natural budget-constrained optimal report*  $\mathbf{q}^*$ , with budget  $b$ , is the choice among all  $\mathbf{q}$  such that  $nb(\bar{\mathbf{q}}, \mathbf{q}) \leq b$ , that maximizes  $\mathbf{p} \cdot \mathbf{S}(\mathbf{q})$ . Let  $\mathbf{q}^*(\mathbf{p}, \bar{\mathbf{q}}, b)$  denote the optimal report with budget  $b$ .

We can now state our desired property:

**Definition 2** *A scoring rule  $\mathbf{S}$  satisfies the budget-constrained truthfulness property if the budget-constrained optimal choice is always a mixture of the initial distribution  $\bar{\mathbf{q}}$  and the belief  $\mathbf{p}$ , and as close to  $\mathbf{p}$  as possible:*

$$\forall \mathbf{p}, \bar{\mathbf{q}}, b > 0 \quad \mathbf{q}^*(\mathbf{p}, \bar{\mathbf{q}}, b) = \alpha \mathbf{p} + (1 - \alpha) \bar{\mathbf{q}}$$

where

$$\alpha \in [0, 1] = \max\{\alpha | b(\bar{\mathbf{q}}, \alpha \mathbf{p} + (1 - \alpha) \bar{\mathbf{q}}) \leq b\}$$

This definition inherently implies that  $\mathbf{S}$  is a proper scoring rule (by taking a sufficiently large  $b$ ); however, it is significantly stronger than the properness condition.

### 3 An impossibility result for unrestricted multi-outcome distributions

For two-outcome events, the budget-constrained truthfulness property holds trivially for all common scoring rules, as the natural budget and expected score both increase monotonically as  $\mathbf{q}$  moves from  $\bar{\mathbf{q}}$  towards the agent's true belief  $\mathbf{p}$ . However, well-known scoring rules do not satisfy this for general multiple-outcome events. Here, we show that this is not accidental: We will prove that no scoring rule (satisfying our technical assumptions) over three or more outcomes can satisfy the budget-constrained truthfulness property. We will assume without loss of generality that there are exactly three outcomes  $X, Y, Z$ ; the result holds *a fortiori* for  $k > 3$ .

We will prove the impossibility result by contradiction. Suppose that we had a scoring rule  $\mathbf{S}$  such that  $\mathbf{S}$  satisfied the budget-constrained truthfulness property.

We first make a simple observation about the local structure of the scoring surface:

**Lemma 1** *Consider a point  $\mathbf{p}$  in the interior of  $\Delta$ , and consider a neighboring point of the form  $\mathbf{p} + \delta \mathbf{\epsilon}$ , where  $\mathbf{\epsilon}$  is a nonzero vector such that  $\epsilon_X + \epsilon_Y + \epsilon_Z = 0$ , and  $\delta$  is an infinitesimally small quantity. Then, there exists a direction  $\mathbf{\epsilon}$  such that  $\frac{dS_X(\mathbf{p} + \delta \mathbf{\epsilon})}{d\delta} = 0$ . Further, for this  $\mathbf{\epsilon}$ ,  $\frac{dS_Y(\mathbf{p} + \delta \mathbf{\epsilon})}{d\delta}$  and  $\frac{dS_Z(\mathbf{p} + \delta \mathbf{\epsilon})}{d\delta}$  must be nonzero and have opposite signs.*

**Proof:** Expanding  $S_X$  around  $\mathbf{p}$  in terms of its partial derivatives, we have:

$$\frac{dS_X(\mathbf{p} + \delta \mathbf{\epsilon})}{d\delta} = \epsilon_X \frac{\partial S_X}{\partial p_X} + \epsilon_Y \frac{\partial S_X}{\partial p_Y} + \epsilon_Z \frac{\partial S_X}{\partial p_Z}$$

Setting the LHS to 0, we get one linear constraint on  $\epsilon_X$  and  $\epsilon_Y$ ; any solution to this will meet our purposes.

For the second part of the statement, consider the point  $\mathbf{q} = \mathbf{p} + \delta \mathbf{\epsilon}$ . We have  $S_X(\mathbf{q}) = S_X(\mathbf{p})$ . However, as this is a proper scoring rule, we must have  $\mathbf{p} \cdot [S(\mathbf{p}) - S(\mathbf{q})] > 0$  and  $\mathbf{q} \cdot [S(\mathbf{p}) - S(\mathbf{q})] < 0$ . As  $\mathbf{p}$  and  $\mathbf{q}$  have non-negative entries, the only way in which this can happen is if  $S(\mathbf{p}) - S(\mathbf{q})$  has one positive and one negative entry. ■

We next establish an intuitive result: given two points  $\mathbf{p}, \bar{\mathbf{q}}$ , there is a budget  $b$  such that the constrained-optimal report is the midpoint between  $\mathbf{p}$  and  $\bar{\mathbf{q}}$ .

**Lemma 2** *Consider a scoring rule  $\mathbf{S}$  that satisfies the budget-constrained truthful property. For any pair  $\mathbf{p}, \bar{\mathbf{q}} \in \Delta$ , there is a  $b$  such that  $\mathbf{q}^*(\mathbf{p}, \bar{\mathbf{q}}, b) = 0.5\bar{\mathbf{q}} + 0.5\mathbf{p}$ .*

**Proof:** Let  $b = \text{nb}(\bar{\mathbf{q}}, 0.5\bar{\mathbf{q}} + 0.5\mathbf{p})$ . We must have  $b > 0$ , because if  $b \leq 0$ , then an agent who believed  $\bar{\mathbf{q}}$  would make at least as much expected profit by reporting  $0.5\bar{\mathbf{q}} + 0.5\mathbf{p}$  instead of being truthful; this would violate the properness of the scoring rule.

Given the definition of the budget-constrained truthful property, we only need to show that, for any  $\alpha > 0.5$ ,  $\text{nb}(\bar{\mathbf{q}}, \alpha\mathbf{p} + (1 - \alpha)\bar{\mathbf{q}}) > b$ . Without loss of generality, let us assume that  $S_X(\bar{\mathbf{q}}) - S_X(0.5\mathbf{p} + 0.5\bar{\mathbf{q}}) = b > 0$ . Given a  $\alpha > 0.5$ , we note that  $(0.5\mathbf{p} + 0.5\bar{\mathbf{q}})$  is a mixture of  $\bar{\mathbf{q}}$  and  $(\alpha\mathbf{p} + (1 - \alpha)\bar{\mathbf{q}})$ . By our quasiconcavity assumption, it must be true that  $S_X(\alpha\mathbf{p} + (1 - \alpha)\bar{\mathbf{q}}) < S_X(0.5\mathbf{p} + 0.5\bar{\mathbf{q}})$ . But then, by the definition of the budget, we must have  $\text{nb}(\bar{\mathbf{q}}, \alpha\mathbf{p} + (1 - \alpha)\bar{\mathbf{q}}) > b$ . ■

The following lemma will allow us to construct a contradiction:

**Lemma 3** *Suppose that  $\mathbf{S}$  is continuous, differentiable, and satisfies the budget-constrained optimality property. Then, there are probability distributions  $\mathbf{p}, \mathbf{q}, \bar{\mathbf{q}}, \mathbf{p}', \bar{\mathbf{q}}'$  and budgets  $b, b'$  such that the following properties are satisfied:*

1.  $\mathbf{q} = 0.5\mathbf{p} + 0.5\bar{\mathbf{q}}$  and  $\mathbf{q} = \mathbf{q}^*(\mathbf{p}, \bar{\mathbf{q}}, b)$ . More specifically,  $S_X(\bar{\mathbf{q}}) - S_X(\mathbf{q}) = b$  and  $S_Y(\bar{\mathbf{q}}) - S_Y(\mathbf{q}) < b$ ,  $S_Z(\bar{\mathbf{q}}) - S_Z(\mathbf{q}) < b$ .

2.  $\mathbf{q} = 0.5\mathbf{p}' + 0.5\bar{\mathbf{q}}'$  and  $\mathbf{q} = \mathbf{q}^*(\mathbf{p}', \bar{\mathbf{q}}', b')$ . More specifically,  $S_X(\bar{\mathbf{q}}') - S_X(\mathbf{q}) = b'$  and  $S_Y(\bar{\mathbf{q}}') - S_Y(\mathbf{q}) < b'$ ,  $S_Z(\bar{\mathbf{q}}') - S_Z(\mathbf{q}) < b'$ .
3.  $\frac{p_Y}{p_Z} \neq \frac{p'_Y}{p'_Z}$

**Proof:** We pick arbitrary  $\mathbf{p}$  and  $\bar{\mathbf{q}}$  in the interior of the space of possible distributions, and let  $\mathbf{q} = 0.5\bar{\mathbf{q}} + 0.5\mathbf{p}$  be a mixture of these two distributions. Let  $b = \text{nb}(\bar{\mathbf{q}}, \mathbf{q})$ . By lemma 2,  $\mathbf{q} = \mathbf{q}^*(\mathbf{p}, \bar{\mathbf{q}}, b)$ . Now, the budget  $b$  is the worst-case loss of moving from  $\bar{\mathbf{q}}$  to  $\mathbf{q}$ ; without loss of generality, assume that this loss occurs with outcome  $X$ . Thus, we must have:

$$S_X(\bar{\mathbf{q}}) - S_X(\mathbf{q}) = b$$

Suppose that we also have  $S_Y(\bar{\mathbf{q}}) - S_Y(\mathbf{q}) = b$ . Then, it must be the case that  $S_Z(\bar{\mathbf{q}}) < S_Z(\mathbf{q})$ . By Lemma 1, we can perturb our choice of  $\mathbf{p}$  slightly such that only  $S_X(\bar{\mathbf{q}}) - S_X(\mathbf{q}) = b$ , while  $S_Y(\bar{\mathbf{q}}) - S_Y(\mathbf{q}) < b$  and  $S_Z(\bar{\mathbf{q}}) - S_Z(\mathbf{q}) < b$ .

Now, consider  $\bar{\mathbf{q}}' = \bar{\mathbf{q}} - \boldsymbol{\varepsilon}$  and  $\mathbf{p}' = \mathbf{p} + \boldsymbol{\varepsilon}$ , where  $\boldsymbol{\varepsilon}$  is a small perturbation that is constrained to ensure that  $\bar{\mathbf{q}}'$  and  $\mathbf{p}'$  are still probability distributions. Note that  $\mathbf{q} = 0.5\bar{\mathbf{q}}' + 0.5\mathbf{p}'$ . By lemma 2, there is some budget  $b'$  such that  $\mathbf{q} = \mathbf{q}^*(\mathbf{p}', \bar{\mathbf{q}}', b')$ . By the continuity of  $\mathbf{S}$ , for all  $\boldsymbol{\varepsilon}$  within a sufficiently small ball, it must still be true that  $S_X(\bar{\mathbf{q}}) - S_X(\mathbf{q}) > S_Y(\bar{\mathbf{q}}) - S_Y(\mathbf{q}), S_Z(\bar{\mathbf{q}}) - S_Z(\mathbf{q})$ .

To get the last part of the construction, we pick  $\boldsymbol{\varepsilon}$  such that  $p'_Y > p_Y, p'_Z < p_Z$ . ■

Finally, we can use the constructed structure of Lemma 3 to prove a contradiction.

**Theorem 4** *There is no scoring function  $\mathbf{S}$  that satisfies the technical conditions (continuity, differentiability, and quasiconcavity) and also satisfies the budget-constrained truthfulness property.*

**Proof:** Suppose that there is such a scoring function. Consider  $\mathbf{p}, \mathbf{q}, \bar{\mathbf{q}}, \mathbf{p}', \bar{\mathbf{q}}'$  and budgets  $b, b'$  as constructed in Lemma 3.

By assumption, if the initial position is  $\bar{\mathbf{q}}$ , and an agent has belief  $\mathbf{p}$  and budget  $b$ , the optimal report is  $\mathbf{q}$ . With this report, the agent's expected score is  $\mathbf{p} \cdot \mathbf{S}(\mathbf{q}) - \mathbf{p} \cdot \mathbf{S}(\bar{\mathbf{q}})$ .

Now, consider the neighborhood of the report  $\mathbf{q}$ . By Lemma 1, there is a direction  $\boldsymbol{\varepsilon}$  such that  $S_X(\mathbf{q} + \boldsymbol{\varepsilon}d\delta) = S_X(\mathbf{q})$ . By lemma 3, and the continuity of  $\mathbf{S}$ , this means that  $\text{nb}(\bar{\mathbf{q}}, \mathbf{q} + \boldsymbol{\varepsilon}d\delta) = \text{nb}(\bar{\mathbf{q}}, \mathbf{q}) = b$ , as locally  $S_X$  is the binding constraint on the budget. In other words, the report  $\mathbf{q} + \boldsymbol{\varepsilon}d\delta$ , as well as  $\mathbf{q} - \boldsymbol{\varepsilon}d\delta$ , are in the feasible set for an agent with budget  $b$ . As  $S_X$  is constant along direction  $\boldsymbol{\varepsilon}$ , the factors affecting the agent's expected score are the change in  $S_Y$  and  $S_Z$ . The first-order conditions for optimality at  $\mathbf{q}$  then give us:

$$p_Y \frac{dS_Y(\mathbf{q} + \boldsymbol{\varepsilon}d\delta)}{d\delta} + p_Z \frac{dS_Z(\mathbf{q} + \boldsymbol{\varepsilon}d\delta)}{d\delta} = 0$$

Rearranging, we must have:

$$\frac{\frac{dS_Z(\mathbf{q} + \boldsymbol{\varepsilon}d\delta)}{d\delta}}{\frac{dS_Y(\mathbf{q} + \boldsymbol{\varepsilon}d\delta)}{d\delta}} = -\frac{p_Y}{p_Z}$$



Note that the left hand side depends only on  $\mathbf{q}$ . Therefore, if we repeat our analysis with  $\mathbf{p}'$ ,  $\bar{\mathbf{q}}'$ , and  $b'$  we would derive that the same LHS (with the same vector  $\mathbf{\epsilon}$ ) is equal to  $-\frac{p'_Y}{p'_Z}$ . By the last assertion in Lemma 3, this is a contradiction. ■

Thus, for unrestricted distributions over 3 or more outcomes, it is impossible to find a smooth scoring rule that will always incentivize agents to move towards their true belief, given the natural budget constraint.

In fact, we can prove an even stronger negative result about the aggregative properties of the market scoring rule in the presence of hard budget constraints: We can show that the optimal report of a budget-constrained agent can be locally insensitive to changes in her belief  $\mathbf{p}$ . This implies that future traders may not be able to infer this trader's belief from her trade, thus resulting in imperfect aggregation of information even if all budgets are known.

In order to prove this result, we first prove two technical lemmas. The first result just shows that there is a reports  $\mathbf{r}$  such that two budget constraints are tight.

**Lemma 5** *Given a strictly proper scoring rule  $\mathbf{S}$  that satisfies the technical assumptions (differentiability and quasi-concavity), and given any starting probability distribution  $\bar{\mathbf{q}}$  in the interior of the space of distributions over  $\{X, Y, Z\}$ , there is a distribution  $\mathbf{r}$  and constants  $a, b > 0$  such that two of the budget constraints with budget  $b$  are tight:*

$$S_X(\mathbf{r}) - S_X(\bar{\mathbf{q}}) = -b; S_Y(\mathbf{r}) - S_Y(\bar{\mathbf{q}}) = -b; S_Z(\mathbf{r}) - S_Z(\bar{\mathbf{q}}) = a$$

**Proof:** Consider the set of all distributions as two-dimensional space. As  $\bar{\mathbf{q}}$  is an interior distribution, there is a circle  $Q = \{\mathbf{q}\}$  of distributions such that  $|\mathbf{q} - \bar{\mathbf{q}}| = c$ . Consider the function  $\text{nb}(\bar{\mathbf{q}}, \mathbf{q})$  on  $Q$ . As  $Q$  is a compact space, and  $\text{nb}$  is continuous, this function must have a minimum value, call it  $2b$ , that is achieved within  $Q$ . As this is a proper scoring rule, we must have  $2b > 0$ .

Now, consider some distribution  $\mathbf{r}'$  such that  $\text{nb}(\bar{\mathbf{q}}, \mathbf{r}') = b$ . Thus,  $\mathbf{r}'$  must lie within the circle  $Q$ . Without loss of generality, let us suppose that  $S_X(\mathbf{r}') - S_X(\bar{\mathbf{q}}) = -b$ . Then, by lemma 1, there is a direction of movement around  $\mathbf{r}'$  such that  $S_X(\cdot)$  is constant, while  $S_Y(\cdot)$  decreases. Moving along this direction, we must reach a point  $\mathbf{r}$  at which both  $S_X(\mathbf{r}) - S_X(\bar{\mathbf{q}}) = -b$  and  $S_Y(\mathbf{r}) - S_Y(\bar{\mathbf{q}}) = -b$ . At this point  $\mathbf{r}$ , we must have  $a \stackrel{\text{def}}{=} S_Z(\mathbf{r}) - S_Z(\bar{\mathbf{q}}) > 0$ , because otherwise  $\bar{\mathbf{q}}$  would earn a higher score than  $\mathbf{r}$  for all three outcomes, which is impossible in a proper scoring rule. ■ Next, we show that for beliefs in the neighborhood of such a point  $\mathbf{r}$ , we can find a 2-dimensional open set with certain properties:

**Lemma 6** *Given the construction of Lemma 5, consider perturbations to the distribution  $\mathbf{r}$  parameterized by a pair  $(\epsilon_X, \epsilon_Y)$ . Define  $\epsilon_Z = -(\epsilon_X + \epsilon_Y)$  and consider distributions  $\mathbf{p} = \mathbf{r} + \mathbf{\epsilon}$ . Then, there is an open neighborhood  $N$  of pairs  $(\epsilon_X, \epsilon_Y)$  (or equivalently, a neighborhood of distributions  $\mathbf{p}$ ) such that the following conditions hold:*

1.  $\epsilon_X < 0, \epsilon_Y < 0, \epsilon_Z > 0$

$$2. S_Z(\mathbf{p}) \geq S_Z(\mathbf{r}) - a$$

$$3. \text{ Either } S_X(\mathbf{p}) < S_X(\mathbf{r}) \text{ or } S_Y(\mathbf{p}) < S_Y(\mathbf{r}), \text{ or both.}$$

**Proof:** Consider the set  $N_1$  of all  $\mathbf{\epsilon}$  such that condition (1) is satisfied, and  $\mathbf{\epsilon}$  is within a small enough ball that  $\mathbf{p}$  is a valid distribution; this set is an open set. Now, consider all  $\mathbf{\epsilon}$  within a small ball such that  $S_Z(\mathbf{p}) \geq S_Z(\mathbf{r}) - a$ . Because the scoring rule is continuous, this set must include an open set  $N_2$  around  $\mathbf{\epsilon} = (0, 0, 0)$ . Now,  $N_1$  and  $N_2$  have a non-empty intersection  $N = N_1 \cap N_2$ ;  $N$  is also an open set.

Thus, it remains to show that, for all  $\mathbf{\epsilon}$  in  $N$ , condition (3) is satisfied. We prove this by contradiction. Suppose that there was a  $\mathbf{\epsilon} \in N$  such that condition (3) was not satisfied, i.e.,  $S_X(\mathbf{p}) \geq S_X(\mathbf{r})$  and  $S_Y(\mathbf{p}) \geq S_Y(\mathbf{r})$ . We must have  $S_Z(\mathbf{p}) < S_Z(\mathbf{r})$ . Then, we would have:

$$\begin{aligned} (\mathbf{p} - \mathbf{r}) \cdot [\mathbf{S}(\mathbf{p}) - \mathbf{S}(\mathbf{r})] &= \mathbf{\epsilon} \cdot [\mathbf{S}(\mathbf{p}) - \mathbf{S}(\mathbf{r})] < 0 \\ \Rightarrow \mathbf{p} \cdot [\mathbf{S}(\mathbf{p}) - \mathbf{S}(\mathbf{r})] &< \mathbf{r} \cdot [\mathbf{S}(\mathbf{p}) - \mathbf{S}(\mathbf{r})] \end{aligned}$$

The first inequality follows from the constructed signs of  $\mathbf{\epsilon}_X, \mathbf{\epsilon}_Y$ , and  $\mathbf{\epsilon}_Z$ . However, for a proper scoring rule, the left hand side of the last inequality is positive, while the right hand side is negative, leading to a contradiction. ■

Now, we can show that, for all beliefs  $\mathbf{p}$  in this 2-dimensional open set  $N$ , the optimal report would be the same,  $\mathbf{r}$ .

**Theorem 7** *Given any proper scoring rule satisfying the technical assumptions, and any interior starting position  $\bar{\mathbf{q}}$ , there is a 2-dimensional open set  $N$  of beliefs  $\mathbf{p}$ , a budget  $b$ , and a feasible report  $\mathbf{r}$  such that, for all beliefs in  $N$ , the optimal budget-constrained report is  $\mathbf{r}$ .*

**Proof:** Construct  $N, b$ , and  $\mathbf{r}$  according to Lemmas 5 and Lemma 6. Consider  $\mathbf{p}$  in  $N$ . We can see, by the third condition in Lemma 6, that  $\mathbf{p}$  is not a feasible report given budget  $b$ .

Now, consider any feasible report  $\mathbf{q} \neq \mathbf{r}$ . As  $\mathbf{q}$  is feasible, we must have  $S_X(\mathbf{q}) \geq S_X(\mathbf{r})$  and  $S_Y(\mathbf{q}) \geq S_Y(\mathbf{r})$ . As  $\mathbf{S}$  is a proper scoring rule, it follows that we must have  $S_Z(\mathbf{q}) < S_Z(\mathbf{r})$ . Now, for  $\mathbf{p}$  and  $\mathbf{\epsilon} = \mathbf{p} - \mathbf{r}$ , we have:

$$\begin{aligned} (\mathbf{p} - \mathbf{r}) \cdot [\mathbf{S}(\mathbf{q}) - \mathbf{S}(\mathbf{r})] &= \mathbf{\epsilon} \cdot [\mathbf{S}(\mathbf{q}) - \mathbf{S}(\mathbf{r})] < 0 \\ \Rightarrow \mathbf{p} \cdot [\mathbf{S}(\mathbf{q}) - \mathbf{S}(\mathbf{r})] &< \mathbf{r} \cdot [\mathbf{S}(\mathbf{q}) - \mathbf{S}(\mathbf{r})] \end{aligned}$$

As  $\mathbf{S}$  is proper, the RHS of the last inequality is negative. Thus, the LHS must also be negative, implying that, under belief  $\mathbf{p}$ ,  $\mathbf{r}$  gives a higher expected score than  $\mathbf{q}$ . As this is true for all feasible  $\mathbf{q}$ ,  $\mathbf{r}$  must be the optimal report. ■

This result shows that failure of information aggregation cannot be avoided with the market scoring rule, *even if all budgets are common knowledge*: For an open set of beliefs, the corresponding report is the same

$\mathbf{r}$ , and hence future traders cannot infer or condition on the precise belief held by a trader. For forecasting problems with  $k > 3$  outcomes, we believe that it should be possible to prove an extension to theorem 7 that shows that the report is insensitive to beliefs in a  $(k - 1)$ -dimensional neighborhood; this is a direction for future work.

## 4 Impossibility results for multi-dimensional events

For some multi-outcome events, we may have reasonable constraints on the family of distributions that agents may believe and report. Of course, there are innumerable different ways of defining constrained distributions over the outcomes. In this section, we consider one very natural form of constraint that arises when the outcome itself can be decomposed into independent dimensions. For example, when eliciting beliefs about the likelihood of a hurricane striking a city this year and next year, it may be natural to assume that, although there are four possible outcomes overall, the likelihood of a strike is independent in the two years, and as such, we can elicit independent probabilities for each year. In this section, however, we show that Theorem 4 extends to this class of restricted preferences as well.

Without loss of generality, we focus on a model with two dimensions of outcome and two possibilities in each dimension. (The impossibility result directly follows for richer outcome and report spaces.) Suppose that the outcome consists of a pair, with the first component either *Top* or *Bottom*, and the second component either *Left* or *Right*. Each forecast  $\mathbf{q}$  is also a pair  $(q_T, q_L)$ . In this model, we assume that the events *Top* and *Left* are independent, so that these two parameters suffice to determine the outcome. We will use the notation  $\hat{\mathbf{q}}$  to denote the vector  $(q_T q_L, q_T(1 - q_L), (1 - q_T)q_L, (1 - q_T)(1 - q_L))$ , i.e., the probabilities of  $(TL, TR, BL, BR)$  implied by a report vector  $\mathbf{q} = (q_T, q_L)$ .

Given the set of pairwise outcomes  $\{TL, TR, BL, BR\}$ , we can represent a scoring rule as a set of four functions:

$$\mathbf{S}(\mathbf{q}) = [S_{TL}(\mathbf{q}), S_{TR}(\mathbf{q}), S_{BL}(\mathbf{q}), S_{BR}(\mathbf{q})]$$

Note that we do *not* assume that the scoring rule is separable into separate scoring rules for each dimension. An agents' score may depend arbitrarily on the pair of outcomes.

Theorem 4 does not immediately imply an impossibility in this domain, because not all distributions over  $\{TL, TR, BL, BR\}$  are expressible in this model. The set of reports is two-dimensional,  $\Delta = \Delta(\{T, B\}) \times \Delta(\{L, R\})$ . (If we had not assumed independence of *Top* and *Left* events, but instead collected a separate probability for each of the four outcomes, then Theorem 4 would immediately apply.) However, we will show below that a very similar construction works in this case as well.

We first prove an analogue of Lemma 1:

**Lemma 8** *Consider a point  $\mathbf{p}$  in the interior of  $\Delta$ , and consider a neighboring point of the form  $\mathbf{p} + \delta \boldsymbol{\epsilon}$ , where  $\boldsymbol{\epsilon}$  is a nonzero vector such that  $(\epsilon_T, \epsilon_L)$ , and  $\delta$  is an infinitesimally small quantity. Then, there exists a direction  $\boldsymbol{\epsilon}$  such that  $\frac{dS_{TL}(\mathbf{p} + \delta \boldsymbol{\epsilon})}{d\delta} = 0$ . Further, for this  $\boldsymbol{\epsilon}$ , out of the three other directional derivatives*

$\frac{dS_{TR}(\mathbf{p}+\delta\mathbf{e})}{d\delta}$ ,  $\frac{dS_{BL}(\mathbf{p}+\delta\mathbf{e})}{d\delta}$ , and  $\frac{dS_{BR}(\mathbf{p}+\delta\mathbf{e})}{d\delta}$ , at least one must be strictly positive, and at least one must be strictly negative.

**Proof:** Expanding  $S_X$  around  $\mathbf{p}$  in terms of its partial derivatives, we have:

$$\frac{dS_X(\mathbf{p}+\delta\mathbf{e})}{d\delta} = \varepsilon_T \frac{\partial S_{TL}}{\partial p_T} + \varepsilon_L \frac{\partial S_{TL}}{\partial p_L}$$

Setting the LHS to 0, we get one linear constraint on  $\varepsilon_T$  and  $\varepsilon_L$ ; any solution to this will meet our purposes.

For the second part of the statement, consider the point  $\mathbf{q} = \mathbf{p} + \mathbf{e}\delta$ . We have  $S_{TL}(\mathbf{q}) = S_{TL}(\mathbf{p})$ . However, as this is a proper scoring rule, we must have  $\hat{\mathbf{p}} \cdot [\mathbf{S}(\mathbf{p}) - \mathbf{S}(\mathbf{q})] > 0$  and  $\hat{\mathbf{q}} \cdot [\mathbf{S}(\mathbf{p}) - \mathbf{S}(\mathbf{q})] < 0$ . As  $\hat{\mathbf{p}}$  and  $\hat{\mathbf{q}}$  have non-negative entries, the only way in which this can happen is if  $\mathbf{S}(\mathbf{p}) - \mathbf{S}(\mathbf{q})$  has at least one strictly positive and at least one strictly negative entry. ■

Lemma 2 does apply in this setting. Therefore, we move on to prove an analogue of Lemma 3:

**Lemma 9** *Suppose that  $\mathbf{S}$  is continuous, differentiable, and satisfies the budget-constrained optimality property. Then, there are probability distributions  $\mathbf{p}, \bar{\mathbf{q}}, \mathbf{q} = 0.5\mathbf{p} + 0.5\bar{\mathbf{q}}$ , a ball radius  $r > 0$  and a budget function  $b(\mathbf{t})$  defined for all  $\mathbf{t} : |\mathbf{t}| < r$  such that the following property is satisfied. Let  $\bar{\mathbf{q}}(\mathbf{t}) = \bar{\mathbf{q}} - \mathbf{t}$  and  $\mathbf{p}(\mathbf{t}) = \mathbf{p} + \mathbf{t}$ . Then,  $\forall \mathbf{t}$  s.t.  $|\mathbf{t}| < r$ ,*

$$\mathbf{q} = \mathbf{q}^*(\mathbf{p}(\mathbf{t}), \bar{\mathbf{q}}(\mathbf{t}), b(\mathbf{t}))$$

and

$$S_{TL}(\bar{\mathbf{q}}(\mathbf{t})) - S_{TL}(\mathbf{q}) = b(\mathbf{t}) > \max\{S_{TR}(\bar{\mathbf{q}}(\mathbf{t})) - S_{TR}(\mathbf{q}), \\ S_{BL}(\bar{\mathbf{q}}(\mathbf{t})) - S_{BL}(\mathbf{q}), S_{BR}(\bar{\mathbf{q}}(\mathbf{t})) - S_{BR}(\mathbf{q})\}$$

**Proof:** We pick arbitrary  $\bar{\mathbf{q}}$  in the interior of the space of possible distributions, and pick a  $\mathbf{q}$  sufficiently close to  $\bar{\mathbf{q}}$  such that  $\mathbf{q} + (\mathbf{q} - \bar{\mathbf{q}})$  and  $\bar{\mathbf{q}} + (\bar{\mathbf{q}} - \mathbf{q})$  are both valid probability distributions. Let  $\mathbf{p} = \mathbf{q} + (\mathbf{q} - \bar{\mathbf{q}})$ , so that  $\mathbf{q} = 0.5\bar{\mathbf{q}} + 0.5\mathbf{p}$ . Let  $b = b(\bar{\mathbf{q}}, \mathbf{q})$ . By lemma 2,  $\mathbf{q} = \mathbf{q}^*(\mathbf{p}, \bar{\mathbf{q}}, b)$ . Now, the budget  $b$  is the worst-case loss of moving from  $\bar{\mathbf{q}}$  to  $\mathbf{q}$ ; without loss of generality, assume that this loss occurs with outcome  $TL$ . Thus, we must have:

$$S_{TL}(\bar{\mathbf{q}}) - S_{TL}(\mathbf{q}) = b$$

Now, consider the other components of the score difference:  $(S_{TR}(\bar{\mathbf{q}}) - S_{TR}(\mathbf{q}), S_{BL}(\bar{\mathbf{q}}) - S_{BL}(\mathbf{q}), S_{BR}(\bar{\mathbf{q}}) - S_{BR}(\mathbf{q}))$ . We first show that we can perturb the choice of  $\mathbf{q}$  (and perhaps permute the outcome names) such that all of these components are strictly less than  $b$ .

Note that at least one of these components must be negative as the scoring rule is proper when the true distribution is  $\mathbf{q}$ . If exactly one of these components is negative, we can swap  $\bar{\mathbf{q}}$  and  $\mathbf{q}$  (changing  $\mathbf{p}$  accordingly), such that there is a unique positive component among all four score differences, which would satisfy the requirement.

Thus, we can restrict our attention to the case in which one component – say  $S_{TR}(\bar{\mathbf{q}}) - S_{TR}(\mathbf{q})$  is positive, while the other two components are negative. In this case, there is a problem only if  $S_{TR}(\bar{\mathbf{q}}) - S_{TR}(\mathbf{q}) = b$ . Now, let us apply Lemma 8. This guarantees a direction of change such that  $S_{TL}$  is unchanged. If  $S_{TR}$  changes along this direction, then we can slightly perturb  $\mathbf{q}$  to get this property. If  $S_{TR}$  is unchanged, then, by lemma 8, both  $S_{BL}$  and  $S_{BR}$  must change, in opposite directions. In the latter case, we can again swap  $\bar{\mathbf{q}}$  and  $\mathbf{q}$ ; this would guarantee that exactly two components are positive, and a small perturbation would guarantee that one of them is uniquely the maximum.

Now, we can safely assume that we have picked  $\bar{\mathbf{q}}$  and  $\mathbf{q}$  such that the only tight budget constraint is that  $S_{TL}(\bar{\mathbf{q}}) - S_{TL}(\mathbf{q}) = b > 0$ . Consider  $\bar{\mathbf{q}}(\mathbf{t}) = \bar{\mathbf{q}} - \mathbf{t}$  and  $\mathbf{p}(\mathbf{t}) = \mathbf{p}' + \mathbf{t}$ , where  $\mathbf{t}$  is a small perturbation. By the continuity of  $\mathbf{S}$ , for all  $\mathbf{t}$  within a ball of sufficiently small radius  $r$ , it must still be true that  $S_{TL}(\bar{\mathbf{q}}) - S_{TL}(\mathbf{q}) > S_{TR}(\bar{\mathbf{q}}) - S_{TR}(\mathbf{q}), S_{BL}(\bar{\mathbf{q}}) - S_{BL}(\mathbf{q}), S_{BR}(\bar{\mathbf{q}}) - S_{BR}(\mathbf{q})$ . Note that  $\mathbf{q} = 0.5\bar{\mathbf{q}}(\mathbf{t}) + 0.5\mathbf{p}(\mathbf{t})$ . By lemma 2, there is some budget  $b(\mathbf{t})$  such that  $\mathbf{q} = \mathbf{q}^*(\mathbf{p}(\mathbf{t}), \bar{\mathbf{q}}(\mathbf{t}), b(\mathbf{t}))$ . ■

Finally, we can use the constructed structure of Lemma 9 to prove a contradiction. This proof is analogous to Theorem 4

**Theorem 10** *For the two-dimensional outcome setting, there is no scoring function  $\mathbf{S}$  that satisfies the technical conditions (continuity, differentiability, and quasiconcavity) and also satisfies the budget-constrained truthfulness property.*

**Proof:** Suppose that there is such a scoring function. Consider  $\mathbf{p}(\mathbf{t}), \mathbf{q}, \bar{\mathbf{q}}(\mathbf{t}), b(\mathbf{t})$  as constructed in Lemma 9.

Fix an arbitrary  $\mathbf{t}$  within the ball, and let us use the shorthand  $\mathbf{p}, \bar{\mathbf{q}}, b$  for  $\mathbf{p}(\mathbf{t}), \bar{\mathbf{q}}(\mathbf{t}), b(\mathbf{t})$  respectively. By assumption, if the initial position is  $\bar{\mathbf{q}}$ , and an agent has belief  $\mathbf{p}$  and budget  $b$ , the optimal report is  $\mathbf{q}$ . With this report, the agent's expected score is  $\hat{\mathbf{p}} \cdot \mathbf{S}(\mathbf{q}) - \hat{\mathbf{p}} \cdot \mathbf{S}(\bar{\mathbf{q}})$ .

Now, consider the neighborhood of the report  $\mathbf{q}$ . By Lemma 8, there is a direction  $\boldsymbol{\epsilon}$  such that  $S_{TL}(\mathbf{q} + \boldsymbol{\epsilon}d\delta) = S_{TL}(\mathbf{q})$ . By lemma 9, and the continuity of  $\mathbf{S}$ , this means that  $b(\bar{\mathbf{q}}, \mathbf{q} + \boldsymbol{\epsilon}d\delta) = b(\bar{\mathbf{q}}, \mathbf{q}) = b$ , as locally  $S_{TL}$  is the binding constraint on the budget. In other words, the report  $\mathbf{q} + \boldsymbol{\epsilon}d\delta$ , as well as  $\mathbf{q} - \boldsymbol{\epsilon}d\delta$ , are in the feasible set for an agent with budget  $b$ . As  $S_{TL}$  is constant along direction  $\boldsymbol{\epsilon}$ , the factors affecting the agent's expected score are the change in  $S_{TR}, S_{BL}$  and  $S_{BR}$ . The first-order conditions for optimality at  $\mathbf{q}$  then give us:

$$\hat{\mathbf{p}} \cdot \left( \frac{dS_{TL}(\mathbf{q} + \boldsymbol{\epsilon}d\delta)}{d\delta} + \frac{dS_{TR}(\mathbf{q} + \boldsymbol{\epsilon}d\delta)}{d\delta} + \frac{dS_{BL}(\mathbf{q} + \boldsymbol{\epsilon}d\delta)}{d\delta} + \frac{dS_{BR}(\mathbf{q} + \boldsymbol{\epsilon}d\delta)}{d\delta} \right) = 0$$

Note that  $\mathbf{q}$  and  $\boldsymbol{\epsilon}$  are independent of our choice of  $\mathbf{t}$ . Thus, for all  $\mathbf{t}$  in an open neighborhood, we must have:

$$\hat{\mathbf{p}}(\mathbf{t}) \cdot \mathbf{D} = 0 \tag{1}$$

where  $\mathbf{D}$  is a vector with first component 0, and at least one positive and one negative component. In other words, all  $\hat{\mathbf{p}}(\mathbf{t})$  corresponding to an open set of reported probabilities  $\mathbf{p}(\mathbf{t})$  must lie in a single linear subspace of 4-dimensional space.

We can now demonstrate a contradiction, by showing that, as the set of feasible  $\hat{\mathbf{p}}$  is curved, no linear subspace can hold an open neighborhood. Consider an arbitrary estimate  $\mathbf{p}_0 = (x, y)$  within this neighborhood, and three points in its neighborhood  $\mathbf{p}_1 = (x + \delta, y)$ ,  $\mathbf{p}_2 = (x, y + \delta)$ , and  $\mathbf{p}_3 = (x + \delta, y + \delta)$ , where  $\delta > 0$ .

Equation 1 must then hold for  $\hat{\mathbf{p}}_0, \hat{\mathbf{p}}_1, \hat{\mathbf{p}}_2, \hat{\mathbf{p}}_3$ . Thus, we must get:

$$\begin{aligned}
(\hat{\mathbf{p}}_1 - \hat{\mathbf{p}}_0) \cdot \mathbf{D} = 0 &\Rightarrow \delta(y, 1 - y, -y, -1 + y) \cdot \mathbf{D} = 0 \\
(\hat{\mathbf{p}}_2 - \hat{\mathbf{p}}_0) \cdot \mathbf{D} = 0 &\Rightarrow \delta(x, -x, 1 - x, -1 + x) \cdot \mathbf{D} = 0 \\
(\hat{\mathbf{p}}_3 - \hat{\mathbf{p}}_1) \cdot \mathbf{D} = 0 &\Rightarrow \delta(x + \delta, -x - \delta, 1 - x - \delta, -1 + x + \delta) \cdot \mathbf{D} = 0 \\
(\text{subtract second eq from third}) &\Rightarrow (1, -1, -1, 1) \cdot \mathbf{D} = 0 \\
(\text{first eq.} - \delta y \times \text{fourth eq.}) &\Rightarrow (0, 1, 0, -1) \cdot \mathbf{D} = 0 \\
(\text{second eq.} - \delta x \times \text{fourth eq.}) &\Rightarrow (0, 0, 1, -1) \cdot \mathbf{D} = 0
\end{aligned}$$

Collecting all the equalities, we have:

$$(1, 0, 0, 0) \cdot \mathbf{D} = 0; (1, -1, -1, 1) \cdot \mathbf{D} = 0; (0, 1, 0, -1) \cdot \mathbf{D} = 0; (0, 0, 1, -1) \cdot \mathbf{D} = 0$$

(The first equality is because, by construction, the first component of  $\mathbf{D}$  is 0.) However, the only solution to this set of equations is  $\mathbf{D} = (0, 0, 0, 0)$ . This contradicts the property, guaranteed by Lemma 8, that  $\mathbf{D}$  has at least two nonzero entries. ■

## 5 A mechanism for budget-constrained elicitation

In this section, we describe a mechanism, the Scaled Scoring Mechanism, for sequential information aggregation that modifies the market scoring rule to address the limitations under budget constraints. The mechanism and its characterization are fairly simple, but it serves to illustrate how designing for budget constraints can be useful. The sequential update and scoring rules we use were previously proposed for binary outcomes in the context of recommender systems by Resnick and Sami [RS07]; that paper did not address the properties for multi-outcome events or the possibility of mis-reporting budgets that we focus on here.

We make one important assumption: Our mechanism will ask users to report their current budget  $b$ , but *we assume that it is impossible for an agent to report a higher budget than her true budget  $b$* , although users may strategically under-report  $b$ . Such a mechanism could be sustained by requiring each agent to deposit her entire reported budget  $b$  up front, at the time of making her report. If the budget is truly a hard constraint, then an agent will not be able to deposit more than  $b$ , even if she is sure that she cannot lose all that she deposits.

Consider a sequence of agents  $1, 2, \dots, m$  interacting with a mechanism, in which the goal of the mechanism designer is to elicit an unrestricted distribution  $\mathbf{q}$  over a given outcome space. Let  $\mathbf{S}$  be a proper scoring rule that satisfies the following properties: Each component of  $\mathbf{S}(\mathbf{q})$  is concave in  $\mathbf{q}$ , and for all allowed  $\mathbf{q}, \bar{\mathbf{q}}$ ,  $\text{nb}(\bar{\mathbf{q}}, \mathbf{q}) \leq B$ . In other words, the maximum budget required for any feasible move is bounded by  $B$ . For example,  $\mathbf{S}$  may be the quadratic scoring rule. If the range of allowed distributions is restricted slightly so that all individual probabilities are bounded away from 0, we could even use the logarithmic scoring rule  $\mathbf{S}(\mathbf{q}) = (\log q_x, \log q_y, \log q_z)$ .

The *Scaled Scoring Mechanism* operates as follows:

1. At any time, the mechanism has a forecast  $\bar{\mathbf{q}}$  that also serves as a reference point. Initialize  $\bar{\mathbf{q}}_0 = q_0$ , an arbitrary initial default prediction.
2. For  $i = 1, 2, \dots, m$ :

(a) Ask agent  $i$  to report a budget (denoted by  $b'_i$ ) and a forecast  $\mathbf{q}_i$ .

(b) Update

$$\bar{\mathbf{q}}_i = \bar{\mathbf{q}}_{i-1} + \max \left\{ 1, \frac{b'_i}{B} \right\} (\mathbf{q}_i - \bar{\mathbf{q}}_{i-1})$$

3. When the outcome is revealed (say, to be  $X$ ), score the agents as follows: each agent  $i$  gets a net payoff equal to:

$$\max \left\{ 1, \frac{b'_i}{B} \right\} [S_X(\mathbf{q}_i) - S_X(\bar{\mathbf{q}}_{i-1})]$$

We make the following observations about the scaled scoring mechanism: Firstly, when reported budgets are unlimited ( $b'_i > B$ ), the updates and scoring are exactly the same as the market scoring rule. However, when budgets are limited, the score is a scaled-down value proportional to the market scoring rule score, *and* the reference point is moved only a fraction of the distance to the reported belief  $\mathbf{q}_i$ .

We can prove the following properties about the Scaled Scoring Mechanism:

**Theorem 11** *The Scaled Scoring Mechanism satisfies the following properties:*

1. *Myopic strategyproofness:* Assuming that each agent  $i$  cannot report  $b'_i$  higher than her true budget  $b_i$ , and that each trader trades just once, it is optimal for each trader to report her true belief  $\mathbf{p}_i$  and her true budget  $b_i$ .
2. *No default:* Under any outcome, each agent  $i$ 's net payoff is at least  $-b'_i$ .
3. *Budget-constrained truthful:* After  $i$  selects her optimal report, the forecast distribution  $\bar{\mathbf{q}}_i$  is a mixture of the earlier forecast distribution  $\bar{\mathbf{q}}_{i-1}$  and  $i$ 's true belief  $\mathbf{p}_i$ , for any budget  $b_i$ .
4. *Limited Loss:* The worst-case loss of the market maker is no worse than the worst-case loss of a market scoring rule market for the same outcome space, with the same initial distribution  $\bar{\mathbf{q}}_0$ .

5. *Myopic sybilproofness: It is never profitable for an agent to divide her budget among multiple identities who make consecutive reports.*<sup>1</sup>

**Proof:** (1): First, we observe that for any fixed reported budget  $b'_i$ , the optimal report  $\mathbf{q}_i$  is equal to agent  $i$ 's belief  $\mathbf{p}_i$ : Her payoff is proportional to  $\mathbf{S}(\mathbf{q}_i) - \mathbf{S}(\bar{\mathbf{q}}_{i-1})$ , and as  $\mathbf{S}$  is a proper scoring rule, her expected payoff is maximized when  $\mathbf{q}_i = \mathbf{p}_i$ . Now, with  $\mathbf{q}_i = \mathbf{p}_i$ , we note that her expected payoff is:

$$\max \left\{ 1, \frac{b'_i}{B} \right\} \mathbf{p}_i \cdot [S(\mathbf{p}_i) - S_X(\bar{\mathbf{q}}_{i-1})]$$

As  $\mathbf{S}$  is a proper scoring rule,  $\mathbf{p}_i \cdot [S(\mathbf{p}_i) - S_X(\bar{\mathbf{q}}_{i-1})]$  is non-negative. Hence, her expected payoff is non-decreasing in  $b'_i$ , and as she cannot exaggerate her budget, it is optimal for her to also report her true budget  $b'_i = b_i$ .

(2): By definition of  $B$ , agent  $i$ 's payoff cannot be less than  $-b'_i$ .

(3): From part (1), the optimal strategy for  $i$  is to truthfully report her true belief  $\mathbf{p}_i$  and her true budget  $b_i$ . As the update to  $\bar{\mathbf{q}}$  is done by mixing with  $i$ 's report, the budget-constrained truthfulness property holds.

(4): Here, we use the fact that each component of the scoring rule  $\mathbf{S}$  is concave. Let  $\lambda = \max 1, \frac{b'_i}{B}$ . By concavity we have, for outcome  $X$ :

$$S_X(\bar{\mathbf{q}}_i) = S_X((1 - \lambda)\bar{\mathbf{q}}_{i-1} + \lambda\mathbf{q}_i) \quad (2)$$

$$\geq (1 - \lambda)S_X(\bar{\mathbf{q}}_{i-1}) + \lambda S_X(\mathbf{q}_i) \quad (3)$$

$$\Rightarrow S_X(\bar{\mathbf{q}}_i) - S_X(\bar{\mathbf{q}}_{i-1}) \geq \lambda[S_X(\mathbf{q}_i) - S_X(\bar{\mathbf{q}}_{i-1})] \quad (4)$$

Similar inequalities hold for other outcomes  $Y, Z, \dots$ . Note that the right hand side of Equation 4 is precisely the payment made to agent  $i$  by the SSM mechanism. The right hand side of the equation can be interpreted as the payment that the MSR mechanism would have made to agent  $i$  if the sequence of reports was  $\bar{\mathbf{q}}_1, \bar{\mathbf{q}}_2, \dots, \bar{\mathbf{q}}_M$ . Thus, the worst-case loss of the SSM mechanism, over all budgets, report sequences, and outcomes, is no more than the worst-case loss of the corresponding MSR mechanism.

(5): This property also follows from equation 4. Suppose the initial market prediction is  $\bar{\mathbf{q}}_{i-1}$ . Consider a set of consecutive reported distributions  $\mathbf{q}_{i1}, \mathbf{q}_{i2}, \dots, \mathbf{q}_{ij}$ , with corresponding budgets  $b_{i1}, b_{i2}, \dots, b_{ij}$ . With this sequence, the final reference value under SSM would be some distribution  $\bar{\mathbf{q}}_{ij}$ . By equation 4, the total payoff for these  $j$  reports would be less than the MSR payoff of a single move to  $\bar{\mathbf{q}}_{ij}$ .

We will argue that an agent  $i$  with the combined budget of all  $j$  agents could make at least as high a profit with a single report  $\mathbf{q}_i$ . Let  $b = \sum_j b_{ij}$ . If  $b \geq B$ , then agent  $i$  could always move to  $\bar{\mathbf{q}}_{ij}$ , thereby earning (under SSM) an amount equal to the MSR payoff of moving to  $\bar{\mathbf{q}}_{ij}$ .

If  $b < B$ , then consider an agent who believes a distribution  $\mathbf{p}_i$ . It suffices to consider the case  $j = 2$ ; higher values of  $j$  will follow by induction, because we can first replace the last two trades with a single

---

<sup>1</sup>This is a relaxation of the path-invariance property characteristic of market scoring rules, in which the sum of payoffs of consecutive trades is equal to the payoff of a single trade with the same final report.



trade without hurting profits, and then repeat this procedure. Let  $u_0, u_1, u_2$  denote the expected scores of  $\bar{\mathbf{q}}_{i-1}, \mathbf{q}_{i1}, \mathbf{q}_{i2}$  under this belief, and let  $v_1$  denote the expected score of  $\bar{\mathbf{q}}_{i1}$  under this belief. For conciseness, let us define  $\lambda_t \stackrel{\text{def}}{=} \frac{b_{it}}{B}$  and  $\bar{\lambda}_t \stackrel{\text{def}}{=} 1 - \lambda_t$ . By equation 4, we see that  $v_1 \geq \bar{\lambda}_1 u_0 + \lambda_1 u_1$ . The total payoff of the 2 trades is seen to be:

$$\begin{aligned}
& \lambda_1[u_1 - u_0] + \lambda_2[u_2 - v_1] \\
& \leq \lambda_1[u_1 - u_0] + \lambda_2[u_2 - \lambda_1 u_1 - \bar{\lambda}_1 u_0] \\
& = \lambda_1[u_1 - u_0] + \lambda_2[u_2 - u_0] - \lambda_2 \lambda_1 [u_1 - u_0] \\
& = \lambda_1 \bar{\lambda}_2 [u_1 - u_0] + \lambda_2 [u_2 - u_0] \\
& \leq (\lambda_1 + \lambda_2) \max\{u_1 - u_0, u_2 - u_0\}
\end{aligned}$$

The right hand side of the last equation is the payoff of a single trade that reported the better of  $\mathbf{q}_{i1}, \mathbf{q}_{i2}$  (under the believed distribution) with budget  $b = b_{i1} + b_{i2}$ . Thus, the single trade never has worse payoff than the two consecutive trades. Inductively replacing two consecutive trades with a single trade, we can show that this result holds for  $j > 2$  as well. ■

We note that both the strategyproofness and resistance to sybil attacks is of a myopic nature, in that it may not hold if agents can make trades before and after honest traders' trades. However, this restriction seems unavoidable, as even in the setting without budget constraints, such non-myopic attacks can be profitable under some models of trader beliefs [CDS<sup>+</sup>10].

When we receive a forecast of  $\mathbf{q}_i$  from agent  $i$ , and know that it is rational for agent  $i$  to report this honestly, it may be tempting to update the market forecast to  $\mathbf{q}_i$  instead of  $\bar{\mathbf{q}}_i$ . However, if we used  $\mathbf{q}_i$  as the reference forecast for the next trade, the market would not satisfy the myopic sybilproofness property described above: A trader may profit from splitting her trade into two successive trades.

Note that, if we had updated the market forecast and reference probability

For each agent  $i$ , if the budget  $b_i > 0$  is known,  $i$ 's true belief can be determined from the updated market forecast  $\bar{\mathbf{q}}$ . (Depending on the form of  $i$ 's signal and the initial  $\bar{\mathbf{q}}_{i-1}$ , this may even be possible if the budgets are not known). Thus, in a model where the budgets are common knowledge, each trader can condition on all past traders' beliefs, as in the budget-unconstrained market. In other words, the SSM market generically enables perfect information aggregation.

The SSM mechanism gets around the impossibility result of section 3 because it enforces a stronger restriction on how an agent may use her budget. In particular, it enforces a form of the *Kelly gambling* [Kel56] strategy, in which an agent invests a fraction of her budget proportional to the magnitude of her unconstrained move. The Kelly gambling strategy is asymptotically optimal for long-term budget growth, but it is suboptimal for expected short-term budget growth. In our model, this manifests itself in the following way: An agent is never able to stake her entire budget in her trade, and consequently, the updated prediction  $\bar{\mathbf{q}}_1$  is

less informative than the updated prediction of the MSR with the same budgets. However, this may be more than offset by the improved ability for future traders to aggregate information, and because the enforced proportional-betting policy leads to better long-term budget growth where agents may be acting myopically.

## 6 Discussion

**Extension to Continuous Outcomes** For geographical events, such as forecasting the location of a tornado strike, it may be natural to model the outcome space as continuous, and set the goal of forecasting the expected value. Lambert *et al.* [LPS08] have shown that (without budget constraints) the mean can be elicited by a proper scoring rule. However, Theorem 10 extends *a fortiori* to a model in which we want to forecast the mean of a 2-dimensional outcome, with independence between the dimensions. A special case of the continuous model is the case in which all beliefs are over the outcomes  $(0,0), (0,1), (1,0), (1,1)$ ; theorem 10 shows that no smooth scoring rule can be budget-constrained truthful in this case, and hence in general.

**Separable scoring rules** For forecasting the mean of a multi-dimensional outcome, one natural approach is to elicit separate mean forecasts along each dimension, and to pay off each dimension's forecast through a separate scoring rule. In other words, we might have a mean latitude market and a mean longitude market. With such a setup, it does not matter whether an agent believes the two dimensions are independent or dependent; her expected payoffs would depend only on her marginal beliefs over the two dimensions. Theorem 10 extends *em a fortiori* to this case as well: For a separable scoring rule, an agent's expected payoff is the same as another agent with the same marginal beliefs who believes the two dimensions are independent, and hence, her behavior will be the same.

**Transformed independence** Because we make no additional assumptions about the structure of the scoring rule in section 4, the results extend to many families of restricted distributions. In particular, any family of restricted beliefs that can be smoothly reparameterized to be independent along two parameters will yield the same result.

**Mechanisms for risk-averse agents** We have modeled traders who are risk-neutral up to a hard budget constraint; this can be viewed as a specific family of risk-averse utility functions. Dimitrov et al. [DSE09] showed that there are no attractive sequential market mechanisms to aggregate information for *all* unknown risk types: any such mechanism must exhibit exponentially reducing payoffs in order to guarantee myopic sybilproofness with a bounded subsidy. The scaled scoring mechanism does not have this negative property. Thus, we have shows one natural class of restricted risk preferences for which it is possible to sequentially aggregate information. We note, however, that this relies on the assumption that you cannot exaggerate your budget; hence, we have access to a mechanism that reveals some information about a trader's risk type.

**Limitations and directions for future work** Our impossibility results required some technical assumptions on the smoothness of the scoring rules. Intuitively, there does not seem to be any essential advantage to using non-smooth scoring rules, and one direction for future work is to relax these assumptions. It would also be helpful to extend theorem 7 to higher-dimensional neighborhoods.

The Scaled Scoring Mechanism relies on the fact that a trader cannot deposit an amount greater than her true budget. However, depending on her reported belief, there may be no outcome in which she loses her entire deposit. Thus, if there is an external credit market from which she can borrow with her trade as collateral, this assumption may be violated. One important question for future work is to find alternative market-like mechanisms for sequential information aggregation in the presence of budget limits.

## Acknowledgments

This work was supported by the National Science Foundation under research grant IIS-0812042 and CCF-0829754. The authors would like to thank Yiling Chen, Nicolas Lambert, John Langford, David Pennock and Daniel Reeves for helpful discussions. Thanks especially to David Pennock and his research group at Yahoo! Research in New York that provided the environment that fostered the research in this paper.

## References

- [All87] Franklin Allen. Discovering personal probabilities when utility functions are unknown. *Management Science*, 33(4):542–544, 1987.
- [CDS<sup>+</sup>10] Yiling Chen, Stanko Dimitrov, Rahul Sami, Daniel Reeves, David Pennock, Robin Hanson, Lance Fortnow, and Rica Gonen. Gaming prediction markets: Equilibrium strategies with a market maker. *Algorithmica*, 58:930–969, 2010.
- [DSE09] Stanko Dimitrov, Rahul Sami, and Marina Epelman. Subsidized prediction markets for risk averse traders. In *Internet and Network Economics*, volume 5929 of *Lecture Notes in Computer Science*, pages 491–497. Springer, Berlin / Heidelberg, 2009.
- [Fri83] D. Friedman. Effective scoring rules for probabilistic forecasts. *Management Science*, 29(4):447–454, 1983.
- [Goo52] I. J. Good. Rational decisions. *Journal of the Royal Statistical Society. Series B (Methodological)*, 14(1):107–114, 1952.
- [Han03] Robin Hanson. Combinatorial information market design. *Information Systems Frontiers*, 5(1):107–119, 2003.

- [Kel56] J. L. Kelly. A new interpretation of information rate. *Bell System Technical Journal*, 35:917–926, 1956.
- [LPS08] Nicolas S. Lambert, David M. Pennock, and Yoav Shoham. Eliciting properties of probability distributions. In *Proceedings of the 9th ACM conference on Electronic commerce*, EC '08, pages 129–138, New York, NY, USA, 2008. ACM.
- [Man04] Charles Manski. Interpreting the prices of prediction markets. Working Paper 10359, National Bureau of Economic Research, 2004.
- [RS07] Paul Resnick and Rahul Sami. The influence limiter: Provably manipulation-resistant recommender systems. In *Proceedings of the 2007 ACM conference on Recommender systems*, RecSys '07, pages 25–32, New York, NY, USA, 2007. ACM.
- [WM68] R. L. Winkler and A. H. Murphy. Good probability assessors. *Journal of Applied Meteorology*, 7:751–758, 1968.
- [WZ06] Justin Wolfers and Eric Zitzewitz. Interpreting prediction market prices as probabilities. Working Paper 12200, National Bureau of Economic Research, 2006.